NASA-CR-203855

FINAL
7N-72-CR
OCIT
021505

DEVELOPMENT OF POINTS AS A PLANETOLOGY INSTRUMENT

GRANT NAGW-2497

Final Report

For the Period 1 October 1990 through 30 November 1995

Principal Investigator

Dr. Robert D. Reasenberg

January 1997

Prepared for
National Aeronautics and Space Administration
Washington, D.C.

Smithsonian Institution
Astrophysical Observatory
Cambridge, Massachusetts 02138

The Smithsonian Astrophysical Observatory
is a member of the
Harvard-Smithsonian Center for
Astrophysics

Final report for NASA grant NAGW-2497
"Development of POINTS as a Planetology Instrument"

NASA support of the POINTS project has ended. This report discusses and points to the results of the final phases of our work.

Most recently, the PI presented an invited paper, "POINTS Mission Studies: Lessons for SIM," at the conference "Planets Beyond the Solar System and the Next Generation of Space Missions" 16-18 October 1996, at the Space Telescope Science Institute. Our paper has been accepted for the proceedings. A copy is appended.

Some of the POINTS work that was nearly finished but not documented at the time that we stopped working on POINTS as a NASA project was the basis for a small project we did for JPL, "Aspects of the POINTS design." The final report to JPL, with copies to several interested people at Code-S, contains four POINTS Technical Memoranda on subjects JPL considered to be applicable to SIM. A few additional copies of this report (of about 150 pages) are available on request.

Additional papers on POINTS (and the related Newcomb instrument), not referenced in our previous reports, include:

R.D. Reasenberg, R.W. Babcock, M.C. Noecker, and J.D. Phillips, "POINTS: The Precision Optical INTerferometer in Space, *Remote Sensing Reviews* (Special Issue Highlighting the Innovative Research Program of NASA/OSSA), Guest Editor: Joseph Alexander, Vol. 8, pp 69-99, Harwood Academic Publishers, 1993. (Invited)

R.D. Reasenberg, R.W. Babcock, M.A. Murison, M.C. Noecker, J.D. Phillips, B.L. Schumaker, and J.S. Ulvestad, "POINTS: an astrometric spacecraft with multifarious applications," in *The Proceedings of the SPIE Conference # 2200 on Space Interferometry*, (Kona, HI, USA, 13-18 March 1994), Vol. 2200, p. 2, 1994. (Invited)

R.D. Reasenberg, R.W. Babcock, M.A. Murison, M.C. Noecker, J.D. Phillips and B.L. Schumaker "POINTS: the instrument and its mission," in the *Proceedings of the SPIE Conference # 2477 on Spaceborne Interferometry II*, (Orlando, FL, April 17-20, 1995), Vol. 2477, p. 167, 1995. (Invited)

J.D. Phillips, R.W. Babcock, M.A. Murison, R.D. Reasenberg, A.J. Bronowicki, M.H. Gran, C.F. Lillie, W. McKinley and R.J. Zielinski, "Newcomb, a small astrometric interferometer," in the *Proceedings of the SPIE Conference # 2477 on Spaceborne Interferometry II*, (Orlando, FL, April 17-20, 1995), Vol. 2477, p. 209, 1995.

J.D. Phillips, "A spectrometer for astronomical interferometry," in the *Proceedings of the SPIE Conference # 2477 on Spaceborne Interferometry II*, (Orlando, FL, April 17-20, 1995), Vol. 2477, p.149, 1995.

M.C. Noecker "Systematic errors in high-precision optical interferometric astrometry," in the *Proceedings of the SPIE Conference # 2477 on Spaceborne Interferometry II*, (Orlando, FL, April 17-20, 1995), Vol. 2477, p. 30, 1995.

R.D. Reasenberg, R.W. Babcock, M.A. Murison, M.C. Noecker, J.D. Phillips, B.L. Schumaker, J.S. Ulvestad, W. McKinley, R.J. Zielinski, and C.F. Lillie, "POINTS: High Astrometric Capacity at Modest Cost *via* Focused Design," in the *Proceedings of the SPIE Conference #2807 on Space Telescopes and Instruments IV*, (Denver, CO, 4-9 August 1996), Vol. 2807, pp 32-50, 1996.

Key Words: exoplanets, astrometry, optical interferometry, mission simulations

POINTS Mission Studies: Lessons for SIM

Robert D. Reasenberg, Robert W. Babcock, John F. Chandler, and James D. Phillips

Smithsonian Astrophysical Observatory, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

Abstract: POINTS (Precision Optical INTerferometer in Space) measures the angle between two widely separated stars. The nominal bright-star measurement accuracy of 2 microarcsec is achieved in two minutes of observing two mag 8 stars. POINTS comprises a metrology system and a pair of independent Michelson stellar interferometers, each with a pair of 35 cm subapertures and a 2 m baseline. The angle between the baselines is adjustable over the range 87 to 93 deg. The POINTS scientific mission is enhanced by a solar shield, which allows observation of stars as close as 10 deg from the Sun. Numerous mission simulations over the past 15 years have elucidated the consequences of the single measurement accuracy and instrument architecture.

For simplicity and efficiency, we divide the target stars into two classes, "reference grid stars" and all others. Grid stars provide reference for other targets, and are also science targets. In the nominal mission, redundant grid-star observations are performed quarterly to determine the stars' positions, proper motions, and parallaxes. We showed more than a decade ago that, if the grid stars are observed with sufficient redundancy, the grid "locks up:" after the observations are combined in a weighted least squares estimate of star positions, proper motions, and parallaxes, the uncertainty in the angle between any pair of grid stars, whether directly observable or not, becomes comparable with the measurement uncertainty.

We have used double-blind Monte-Carlo mission simulations to study the planet-finding capabilities of POINTS and to determine the reliable detection threshold with a nominal observing program. If we demand a negligible probability of false alarms, then with our standard observing schedule, the detection threshold for short-period planets is a signature with amplitude equal to the single-measurement observing precision, and orbital elements can usually be determined. For planets with periods longer than the mission, the threshold rises steeply with period.

These studies of the POINTS mission yield seven lessons for the Space Interferometry Mission (SIM), which are discussed.

1. Introduction

The POINTS (Precision Optical INTerferometer in Space) project started as a casual investigation in 1978 based on a concept originally suggested to NASA in 1976 by Irwin Shapiro (1978). It proceeded at a low level for many years until modest funding was obtained during the 1980s. Because global microarcsecond astrometry rests upon several *terrae incognitae*, the corresponding investigations had to be approached *ab initio*. Among these was a reference frame suitable for microarcsecond measurements. The traditional approach to a reference frame, which is central to astrometry, is to use stable objects in a local field. The POINTS-global approach is to build a "reference grid of stars" through frequent redundant intra-grid observations, and connect the grid to a stable reference such as the quasars. We demonstrated the phenomenon of grid lock-up and the associated instrument-parameter estimation in 1980-81. This formed an essential part of the mission's foundation. In the early 1990's, when we investigated a much simplified astrometric interferometer, Newcomb, finding a way of locking up the grid was among the first and most important of our investigations (Reasenberg et al. 1993). Both the POINTS and Newcomb studies showed that good sky coverage is essential to grid lockup.

During the 1990's, we investigated the ability of POINTS to detect planets around the grid stars. We know of no way to do this reliably except by the kinds of simulations we describe below. These studies confirmed our earlier hope that grid stars could function as reference objects, even when many have substantial motion due to unseen companions. They also led to the realization that the mission's sensitivity to the discovery of planets around distant stars falls sharply with planetary period when that period is longer than about 3/4 of the mission length. This phenomenon places a premium on mission longevity and thus mandates high reliability design.

In Section 2, we discuss the architecture and some mission characteristics of POINTS and Newcomb. We introduce the grid-lockup phenomenon and its breakdown at Sun-exclusion angles of about 90 deg. Section 3 contains the motivation and description of the double-blind simulations of the POINTS planet-finding effort. It is shown that, with a nominal observing schedule, the detection threshold for short-period planets is a signature about equal to the single-measurement accuracy. A simplified system for simulating the detection of planets is introduced in Section 4, where we discuss its calibration against the more labor-intensive approach of Section 3. The work described in Section 4 is reduced in Section 5 to a simple relationship among measurement accuracy, mission length, and target characteristics for 50% probability of detection. Finally, in Section 6, the results are summarized with accent on those most applicable to SIM. Many of the subjects discussed herein are elaborated further in a forthcoming paper by Babcock et al. (1997).

2. The POINTS and Newcomb Interferometers, Instrument and Mission Characteristics

POINTS (Reasenberg et al. 1988, 1993, 1995) measures the angle between two widely separated stars. It comprises a metrology system and a pair of independent Michelson stellar interferometers, each with a pair of 35 cm subapertures and a 2 m baseline. The angle between the baselines is adjustable over the range 87 to 93 deg. The wide separation between simultaneously observed stars makes POINTS a global astrometric instrument and provides three key advantages: (1) reference stars can be anywhere in a 6×360 deg band (5% of the sky). Such a band may be expected to include about 80 stars as bright as visual mag 5 and 2000 stars as bright as mag 8. Integration time is not prolonged by the need to use the faint reference stars that would be expected in a small field. (2) Measurements over the entire sky allow recalibration and bias estimation through 360 deg closure over a time scale of hours for an agile spacecraft free of severe pointing restrictions. (3) Parallax determinations are absolute. There is no need to use "zero parallax objects" which, given the measurement accuracy, would be extragalactic and thus faint. A mission might use 300 grid stars for redundancy since some stars may eventually be shown to have properties that make them unsuitable for precision astrometry. For the grid stars (mag$\leq$8), POINTS would make about 360 measurements per day and complete a set of grid measurements in 4.2 days. For fainter objects, measurements would take longer or have greater error. The limiting magnitude, which is set by sky background, is 14 without a spectrometer slit or 18 with one.

For thermal stability and observational flexibility, the spacecraft is permanently shaded from the Sun by a shield that also supports a solar power array. The shield sees continuous sunlight except for seasonal eclipses of up to two hours duration. A shield of 4.8 m diameter held 4.35 m from the center of the spacecraft would allow observation of stars as close to the Sun as 10 deg. However, near-Sun observations are not critical for planet finding and might pose a hazard to the detectors in the event of a pointing-system failure. Therefore, all simulations were done with a simple 30 deg Sun-exclusion angle.

For simplicity and efficiency, we divide the target stars into two classes, "reference grid stars" and all others. Grid stars provide reference for other (generally less frequently observed) targets, and are also science targets. Our observing strategy begins with the selection of a set of stars for the reference grid. Grid stars should be bright to minimize acquisition and integration time. We have shown that a grid of stars, with none known to be in a multi-star system, can be selected from nearby stars (within 22 pc) of mag 8 or brighter.

In the nominal mission, a redundant set of grid star observations is performed quarterly to determine each star's five astrometric parameters: (2) positions, (2) proper motions, and (1) parallax. If the grid stars are observed with sufficient redundancy, the grid "locks up;" after the observations are combined in a weighted-least-squares estimate of the astrometric parameters of each star, the uncertainty in the angle between any grid pair, whether directly observable or not, becomes of the order of the measurement uncertainty (Chandler & Reasenberg 1990). Grid behavior is largely characterized by a single parameter M, the ratio of the total number of observations that are possible (ignoring obscuration) to the number of grid stars. The grid locks for $M \gtrsim 3.5$, and with increasing $\underline{M}$ the mean inter-star angle uncertainty decreases at a rate that changes from $1/M$ to $1/\sqrt{M}$ as M increases from ~5 to ~20. At $M \approx 4.2$, the modal uncertainty in star pair separation is equal to the single-measurement uncertainty. An overall rotational degeneracy remains, which can only be broken by ties to an inertial reference frame. Quasars are the obvious source for a frame tie; stellar aberration provides a weaker tie to the Earth's orbit.

POINTS is expected to be able to make about 360 measurements per day of targets as bright as the grid stars, so the 1500 quarterly grid measurements can be completed in 4.2 days. The remaining time would be used for observing other targets against the grid, either planet candidates, quasars for the frame tie, or targets of astrophysical interest (Reasenberg 1984, Reasenberg et al. 1988, Working Papers 1991, Peterson et al. 1996). The nominal mission lasts for 10 years; all designs (and cost estimates) were based on that nominal life. After 10 years of quarterly 2 microarcsecond (μas) observations with M=5, grid star position, parallax, and annual proper motion are determined on average to 0.24, 0.16, and 0.08 μas respectively (Reasenberg and Shapiro 1986, Reasenberg

1993).

The accuracy of a single measurement is limited by the number of collected photons and by systematic errors. For our simulations, we use the nominal POINTS accuracy of 2 μas (Reasenberg 1996). Uncertainties and sensitivities derived from the simulations are proportional to this single-measurement accuracy. Control of systematic error is crucial; the entire 2 μas error budget would be consumed by an uncorrected 20 picometer (pm) change in the optical path difference (OPD) between the two sides of one interferometer caused, for example, by displacement of a primary mirror. This subject has been discussed extensively. (See Noecker 1995 and references therein.)

Newcomb We briefly consider Newcomb, which is a smaller, simplified variant of POINTS. Newcomb comprises five interferometers stacked one above the other. Each interferometer's baseline and optical axis is parallel to the instrument's *principal plane*. The second, third, and fourth axes are separated from the first by fixed *observation angles* of 40.91, 60.51, and 70.77 deg. The fifth axis is redundant, parallel to the first. The baseline length is 35 cm and the aperture is 5 cm. Each interferometer detects a dispersed fringe (channeled spectrum), which falls on a short CCD detector array. The optical passband is from 0.9 to 0.3 microns.

To reduce cost, we replaced the POINTS-type articulation with a ±0.35 deg field of view made possible by moving the beamsplitter assembly along the baseline direction. With this field of view, building a POINTS-type grid with randomly placed stars would require about 8000 measurements per (quarterly) observation series of 1600 stars. With the nominal observation time of five minutes (plus slew time), this would be unacceptably slow: too much of the observing time would be consumed measuring the grid, and closure information would be extracted too slowly. Instead, we chose to designate 180 *berths* in a systematic array on the sky, and showed that this pattern could be placed such that a large fraction of the berths would contain bright stars. Each berth is a circular patch of sky of 0.35 deg radius in which we hope to have a grid star. To construct the array of berths, we start with the 60 points that are vertices of the regular truncated icosahedron. The observation angles listed above are inter-vertex angles such that for each vertex there are four other vertices at each chosen angle. (There is no angle offering higher multiplicity.) To the original set of 60 points, we add two additional such sets by rotating the figure ±20.82 deg.

With a grid of stars filling all 180 berths, we showed POINTS-like grid lockup even with large Sun-exclusion angles. In the covariance studies, we assumed nine quarterly observation series and five bias parameters per series per observation angle; we estimated all five astrometric parameters for each star. In an extension of the study, we eliminated stars at random from the set and found that there was stable behavior with as few as half the stars in the grid. Further, reducing from 180 stars in the grid to 120 increased the statistical uncertainty of star position by only 31% due to degeneracy (*i.e.*, with constant number of observations).

With the above as background, we can address the result of a study of the effect of the Sun-exclusion angle on lockup. Figure 1 shows that the uncertainty in grid star positions increases slowly with the Sun-exclusion angle until 90 deg, beyond which angle stars near the ecliptic poles are never observable and the grid starts to unlock. The results shown are for a study that started with 120 of 180 berths filled before applying the Sun-exclusion constraint, but other studies show that a complete grid manifests similar behavior, as do POINTS grids. We conclude that, for the timely production of

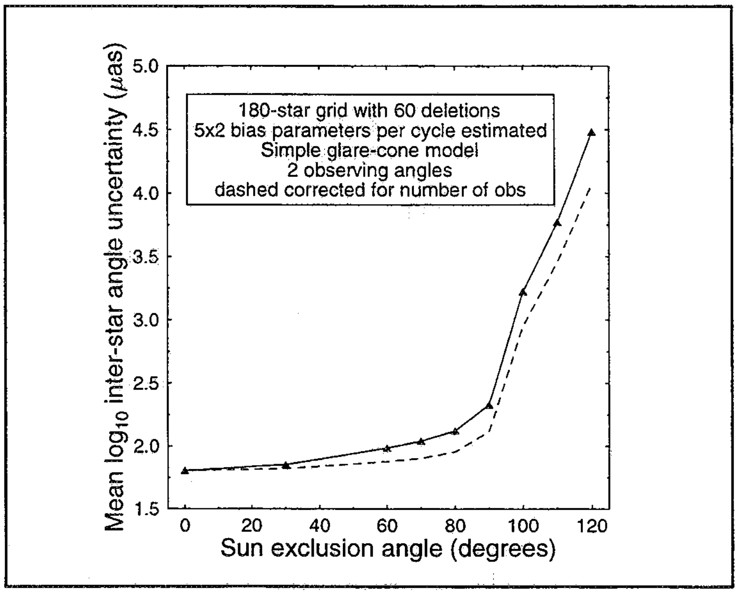

Figure 1 Angular uncertainty *vs.* Sun-exclusion angle for Newcomb. The abscissa is the minimum permissible angle between an observed star and the Sun. For the dashed curve, the observations lost due to the glare-angle constraint are replaced by permissible observations. (Reproduced from Phillips et al. 1995.)

a robust grid, the Sun-exclusion angle should be well under 90 deg.

3. Double-blind Simulations of a Planet-Finding Mission

In a classical astrometric study, there is a target star and a set of reference stars that are normally expected to be more stable than the target. This description applies to POINTS non-grid targets. For POINTS grid stars, the situation is different: there is no distinction between target and reference stars. Further, in our studies we assume that all of the grid stars are likely to have planets. Thus, simulations were needed to show that the grid-star data could be disentangled and to determine the planet detection threshold, the ability to estimate orbital parameters, and the robustness against false alarms. For any global astrometric instrument operating with sufficient accuracy to detect remote planets, there will be a similar need.

Planets orbiting non-grid stars are conceptually easier to detect and characterize because the reference star positions are well known from the grid analysis. Eventually, an optimal estimator would use all available data and the distinction between grid and non-grid stars would be blurred. This approach would be appropriate, for example, at the end of the mission when the problem is linear, the parameter set is fixed, and the need is to make the final adjustments to the parameters. However, it would be an unnecessary complication early in the mission.

Here we describe the techniques that we developed over the past several years to simulate mission data and to investigate the suitability of such data to identify and characterize planets around stars. An unsurprising result from our early simulations is that knowing which stars have planets makes it easier to find them. We therefore adopted a "double-blind" approach to our simulations. The techniques are labor intensive because they address a nonlinear system identification and estimation problem near the noise threshold; practiced judgement is required and automation of the process was considered infeasible during the study. In the next section, we describe a simplified method that we have shown to approximate the behavior of complete simulations, and that requires considerably less human effort to yield results.

The double-blind simulations have three stages: create a fictitious universe that includes a set of randomly generated planets, which we will call the Monte Carlo Realization (MCR), and a corresponding set of "pseudo data;" analyze the pseudo data to yield the Estimated Model (EM) of the location and characteristics of the planets; compare the EM with the MCR. If a simulated analysis is completely successful, the MCR and EM will have planets around the same stars, although the corresponding planet orbits will differ slightly due to simulated measurement noise. However, the process is iterative and, at intermediate iterations, the EM may miss some planets and include some spurious ones; the estimated orbital elements may be far from correct.

In our simulations, we generated planets with random orbital semimajor axis, eccentricity and mass or log(mass), all uniformly distributed over ranges appropriate for testing instrument sensitivity. The orbital angles were generated to yield a uniform distribution of orientations. We generally made the distribution of eccentricities uniform over the range from 0 to 0.2, the range (although not the distribution) found for planets in the solar system. Most (but not all) of our simulations have been limited to at most one planet per star. In the solar system, Jupiter dominates; the contribution from Saturn would be hardly distinguishable from proper motion with observations over 10 years. We typically allowed 20 to 50 planets in a single simulation, all orbiting the 100 grid stars. (A single simulation with 90 massive planets, modeling only circular orbits to limit the number of adjustable parameters, successfully detected every planet. All other simulations had under $\approx$50 planets.)

Our software generated pseudo data comprising periodic (quarterly in these studies) sets of measurements of angles between grid stars, corrupted by white zero-mean Gaussian measurement noise with 2 μas standard deviation. Simulated data were analyzed in the same way that mission data would be analyzed, except that for the latter, it would be necessary to "condition" the data, *i.e.*, to apply corrections derived from spacecraft auxiliary and engineering data and to identify and delete "defective" observations. The model used in our simulations includes star positions, parallaxes, and proper motions; the Earth orbit (needed for parallax, assumed known, and represented by elliptical elements); instrument bias parameters; and perturbations by exoplanets.

All simulations in this section used a double-blind protocol based on Monte Carlo techniques: a colleague (JFC) of the analyst (RWB) would prepare an auxiliary dataset with the seed for the random number generator and with ranges for the allowed mass, radius, eccentricity, and number of planets. The Monte Carlo software would generate the realization but not disclose it. The software would also generate pseudo data and make these available to the analyst. The analyst would then perform a series of numerical experiments with the data until he felt that he had identified all the planets that could be detected reliably. At that time, he would enable the printing of the actual planet parameters for comparison.

Seven double-blind simulations were run of ten-year missions with quarterly observation series at M=5. In these simulations, 205 of 236 planets were detected and

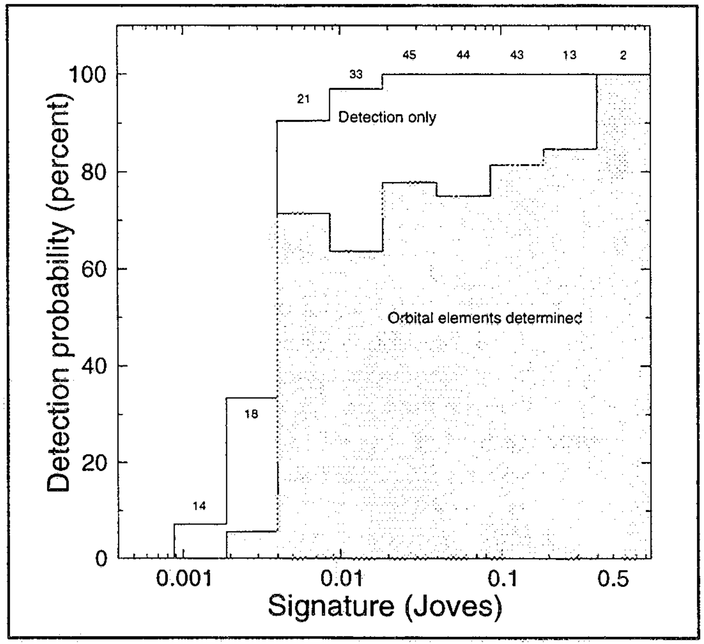

Figure 2. Detection probability from seven double-blind simulations. Numbers over the histogram are the total number of planets generated in each bin. Three undetected planets are off scale to the left.

there were no false alarms. The procedure is robust with the decision-making rules used by the analyst; of the 31 planets not detected, 28 were below the detection threshold established by the study. Results of this phase of the investigation are summarized in Fig. 2 by a histogram of the probability of detection as a function of signature. Here we introduce the "Jove" as a unit of astrometric signature equal to that induced by Jupiter around the Sun as viewed from 10 pc, $\approx$500 μas amplitude.

In some cases, planets could be detected but orbits could not be fully determined: either the eccentricity did not converge to a physical value or the orbit radius grew without bound on successive iterations. In these cases, the analyst would first constrain the model orbit to be circular. If the candidate remained troublesome, the semi-major axis could be constrained to the value determined by the candidate-identification algorithm. The analyst would attempt to lift these constraints as the simulation approached completion, but some orbits continued to require constraints. Almost all of the problem orbits had periods of 8 or more years combined with either edge-on observing geometry or initial parameter guesses too far from correct to allow convergence in a linear estimator. With the nominal observing schedule, the probability of detection is high for signatures at least as large as 0.004 Jove, which happens to be the single-measurement precision, and drops rapidly for smaller signatures. In reporting results of such simulations, it is important to distinguish between detection and determination of planetary orbits.

False alarms Another important consideration is the demand for reliable detections. One can conveniently express confidence limits in terms of the corresponding number of (single parameter) standard deviations. For a planet search, one looks at a large number of possible "detection events," I = (number of stars) × (number of independent planetary periods considered) × (two phases for each period). If we wish the experiment as a whole to have a reliability corresponding to A standard deviations, then the detection threshold for each event must be set to B standard deviations, where the following approximation (see Appendix) may be used for $A > 2$:

$$B \approx \sqrt{A^2 + 2 \log_e I} \tag{1}$$

For our study, which was done conservatively, $I = 7\times100\times20\times2 = 28{,}000$ and we assume A=3. Thus, we obtain $B\approx5.43$; "three sigma" mission reliability requires five+ sigma detection of individual signatures. This does not address the deviation of real noise from Gaussian (the tails are always thick) or noise correlations, as would be expected from some types of astrophysical noise (starspots) and from some of the instrumental errors other than the dominant starlight photon counting statistics[1]. Note that for a mission with 100 (10,000) targets, I=4000 (400,000), and Eq. 1 yields B = 5.1 (5.9) for A = 3.

For the series of seven double-blind simulations, the detection threshold was found *a posteriori* to have been 6.9σ. This (excessively) high value, which by Eq. 1 corresponds to A = 5.2, is related to the three planets in the simulation that should have been detected but were not. It is also consistent with there having been no false alarms in a study of 700 stars. Had resources permitted, we would have done additional simulations with a more aggressive threshold. Any serious discussion of the threshold for the detection of remote planets by a proposed technique should be for the case of a low to negligible level of false alarms.

4. A Simplified System for Planet-Finding Simulations

The simulations described in the previous section have shown that a POINTS mission could find a class of planets, if they exist, and do so without excess false alarms. However, these simulations are time consuming, so it is hard to generate good statistics. In early 1994, while we were doing the simulations, one of us (RWB) hypothesized that the following two statements were nearly the same: (1) Using a standard observing sequence, a planet around the observed star would be detected and its orbital elements could be determined; and (2) Starting near the correct answer, an iterated weighted-least-squares estimator seeking only the orbital elements and mass of the planet and the astrometric parameters of its star would converge. We anticipated that, depending on the details of the algorithms associated with the two statements (for example, the number of reference stars used), a scale factor might need to be applied to one of the planet

---

[1] Note also that the problem is being treated as one-dimensional, i.e., this is the statistics of planet mass. Properly, one should look at the mapping of the noise onto the phase space of orbital elements, and determine the volume of that space in which noise would be interpreted as a planet. For example, false planets of negative mass (but opposite phase) should arise from noise as often as those of positive mass, resulting in a factor two more false detections. This would increase the required detection threshold only a tiny amount, however, because of the exponential nature of the Gaussian distribution.

signatures. With reasonable algorithms, we expect the scale factor to be of order unity. At the PBSS meeting, one of us (RDR) learned that the same hypothesis (Casertano 1996) had been independently developed and used as the basis for the paper by Lattanzi et al. (1997).

Here we discuss an implementation of the above hypothesis, an automated scheme that considers only a single target and its reference stars. We used the more labor intensive mission simulations described in the previous section to provide "ground truth," and found that the simplified system could be calibrated to provide a good approximation to the full simulation. In this section, we describe the simplified system and examine its behavior in the regime where detection probability is falling from near one to near zero. This leads, in Section 5, to a parametric description of POINTS planet-detection capabilities.

In the simplified simulations, seven solar-mass stars were spaced equally around the equator and an eighth was placed near the pole. In the model, all eight stars were mag 8 and at 10 pc; a single planet with random initial conditions orbited the pole star. The planetary system is observed against seven reference stars rather than the average of ten in the full simulations, but the reference stars are equally spaced so chance alignments cannot make one coordinate poorly determined. Quarterly observations were simulated for a 10 year mission.

Starting from the MCR parameter values, which would be right answers absent the effect of the simulated measurement noise, the solutions were iterated up to ten times; the iterations were terminated if convergence was detected or if unphysical elements or divergence were encountered. The adjustable parameters were the mass and orbital elements for the planet and the five astrometric parameters of the pole star. (The star mass was assumed known.) System parameters and solution status (converged or failed) were tabulated for statistical analysis. The parameters that were found to most strongly affect convergence are, not surprisingly, signature amplitude and period. Planet mass and orbital radius are closely coupled to these.

An attempt was made to duplicate the double-blind simulations using the eight-star system. A series of 10-year simulations was made with planet parameters and masses of the polar star varied to match the distributions used in the combined double-blind simulations. The points in Fig. 3 are convergence probabilities in the eight-star system after the application of a calibration factor discussed below. The horizontal error bars represent the binning width, which was chosen to match that used for the full simulations. It is the shaded portion of the histogram in Fig. 2 that should be compared with results in the eight-star system.

A convenient measure of sensitivity is $S_{50}(P,t)$, the signature needed to give a 50% chance of detecting and characterizing a planet with orbital period P in a mission of length t with the nominal observing sequence. Two effects increase S_{50} determined automatically in the eight-star system over that determined by an analyst in the nominal M=5 POINTS reference grid. (1) The pole star is observed against seven reference stars; on average, ten reference stars are available in the nominal grid. (2) Near the sensitivity limit, the best-fit orbital elements may differ significantly from the true elements due to measurement noise. Thus, a solution starting from the correct noise-free elements can be outside the linear regime, and thus may diverge. In manual solutions, the initial elements are derived from the noise-corrupted data, and the analyst has the option of fixing selected orbital elements while attempting to bring the remaining elements into the linear regime, so it is possible to recover from divergences caused by poor starting

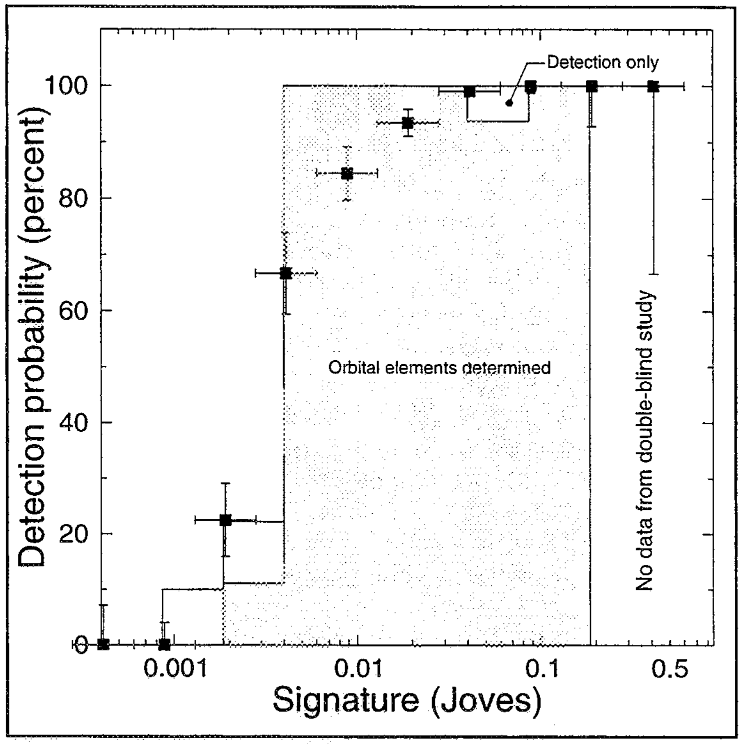

Figure 3. Detection probability from double-blind simulations (histogram) compared with that from the simplified eight-star simulations (points). Double-blind data are those from Fig. 2 having periods of 8 years and less, matching the period distribution used in the eight-star simulations. In performing the eight-star simulations, we attempted to match the star and planet parameters of the full simulation. The points shown are calibrated (shifted in signature) to match the shaded histogram. Vertical error bars are statistical (binomial distribution) and horizontal bars represent binning width.

values. A scale factor of 0.7 was applied to the signatures in the eight-star system to align the points in Fig. 2 with the probability of determining orbital elements found from the double-blind simulations.

5. Mission Length Considerations

We hypothesize that, for a given distribution of eccentricities, S_{50} should have the form

$$S_{50}(P,t) = \sqrt{N_0/Nt}\, f(t/P)\, (\sigma_0/2\,\mu\text{as}) \qquad (2)$$

where the function $f(t/P)$ could be determined empirically in the simple eight-star system by running simulations with a wide range of signatures. In the equation, N_0 is the number of observations per unit time with the nominal observing schedule, N is the actual

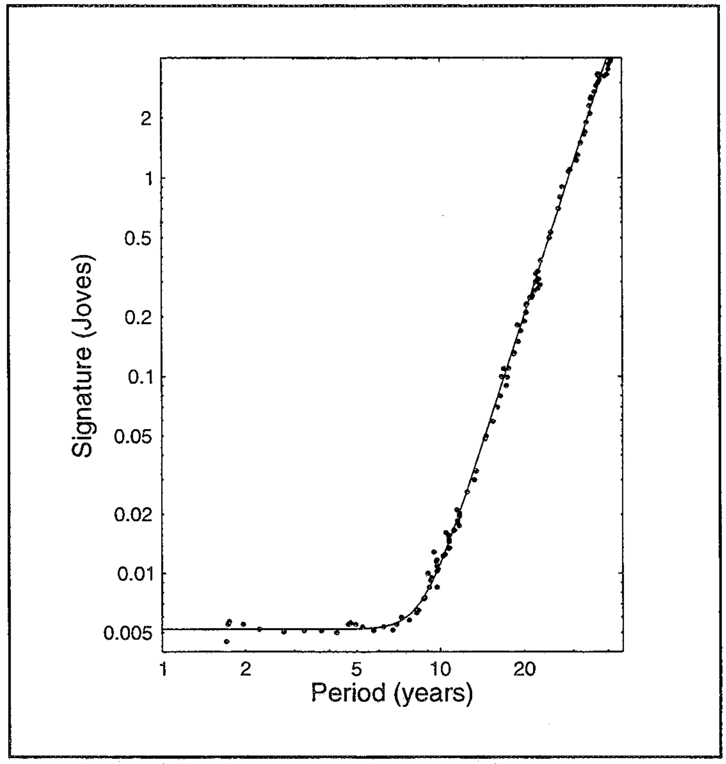

Figure 4. Empirical fit to the signature for 50% detection probability *vs.* period. The points are taken from three simulation series, collectively including 300,000 cases. The solid line is a fit to an RSS of a power law and a constant, see Eq. (3).

number of observations per unit time, and σ_0 is the single measurement precision. The $\sqrt{1/t}$ dependence comes from the increase in the number of measurements with time (ignoring the discretization of quarterly observation cycles).

We performed three series of simulations, each with 100,000 cases. For these, the planet orbital periods spanned 2-12, 2-30, and 2-45 years, and e was uniformly distributed from 0 to 0.2. The corresponding maximum astrometric signatures were 0.02, 0.4, and 4.0 Jove, respectively. From contour plots of the convergence probability as a function of period and signature, we found period-signature pairs for a convergence probability of 50%. In Fig. 4, these points are shown with the correction factor of 0.7 applied. The asymptotic behavior of these points appeared linear (on a log-log plot) for both small and large periods, which gave rise to this form for $f(t/P)$

$$S_{50} = \sqrt{10/t}\sqrt{N_0/N}\sqrt{A^2+\left[B(P/t)^C\right]^2}\,(\sigma_0/2\,\mu\text{as})\ \text{Joves} \tag{3}$$

Using these points, we fit the free parameters of Eq. (3) to obtain A=0.0036, B=0.0069, and C=4.35. The A and B terms in S_{50} are equal at P=0.86t; sensitivity is highest when more than a full orbit is observed and falls rapidly for periods longer than the observation span. The t-dependence of Eq. (3) was confirmed by a selection of simulated missions with durations from 2 to 20 years.

Figure 5 shows S_{50} based on Eq. 3, scaled to mission lengths ranging from 2 to 20 years. However, the assumption of continuous observations is questionable for the shorter missions in this range. Perhaps more interesting than the detectable signature is the detectable mass. The signature in Joves due to a planet of mass M_p and semi-major axis a orbiting a solar-mass star a distance d from Earth is

$$\begin{aligned} S &= \frac{M_p}{M_J}\frac{a}{a_J}\frac{10\,\mathrm{pc}}{d} \\ &\approx \frac{M_p}{M_J}\frac{P^{2/3}}{5.2}\frac{10\,\mathrm{pc}}{d} \end{aligned} \tag{4}$$

where a_J is Jupiter's semi-major axis, approximately 5.2 AU. Mass contours are included in Fig. 5 as an aid in interpreting the meaning of signature and period. It is apparent from the figure that long mission life is central to the successful detection of a variety of planets. The need for long life is even more important when one attempts to detect multiple planets, particularly when there is "frequency crowding."

6. Conclusion

Global microarcsecond astrometry, which must be done from outside Earth's atmosphere, has aspects unlike its ground-based or narrow-field counterparts. In particular, the reference frame can advantageously be constructed by highly redundant observations among a small group of bright "grid" stars. When the grid is rigid, it can be tied to a set of distant stable objects, with quasars as prime candidates. That set can be (but need not be) small, since the positional information is well transferred around the sky. Necessarily, one must study the possibility that the reference objects will themselves have significant motion (of their centers of light). We have investigated aspects of the reference frame, demonstrating grid lock-up and the analysis of the data set for planets. In the process, we have developed a methodology that can be applied to any pointed global astrometric mission. The methodology involved a series of simulations to investigate the detection of planets. Some aspects of this work are specific to POINTS. However, much of it is directly applicable to SIM. Below, we summarize the latter aspects as a series of lessons:

1. For the timely production of a robust grid, the Sun-exclusion angle should be well under 90 deg.

2. For any global astrometric instrument operating with sufficient accuracy to detect remote planets, it will be necessary to demonstrate how and whether it will be possible to disentangle the data. This will be especially important if a large fraction of the stars have planets.

3. Our simulations show that stars with companions or potentially having companions may serve as reference grid stars. When half of the stars each have a single significant planet, there is little loss to the astrometric value of the grid. This has only been shown in the POINTS case, with low Sun-exclusion angle, good temporal coverage, high

redundancy of the grid observations, a wide angle between the observed stars, and an agile spacecraft.

4. An unsurprising result from our early simulations is that, in a mission simulation with manual intervention, knowing which stars have planets makes it easier to find them. A "double-blind" approach to simulation is the only method we know to obtain a reliable value for the detection threshold. (The same software can eventually be used to train the people who will work with the real data.)

5. In reporting results of a mission simulation, it is important to distinguish between detecting a planet and determining its mass and orbital elements. Beyond that is the question of the accuracy with which those parameters are determined. Further, any serious discussion of the threshold for the detection of remote planets by a proposed technique should be for the case of a low to negligible level of false alarms, and that level should be estimated and included in the discussion.

6. A substantial savings in computation time can be had by examining convergence of the iterated solution, rather than performing a complete analysis of simulated missions. However, the complete analysis is required to establish the reliability and scaling of the results from the "convergence test" approach.

7. The range of planets that can be detected is highly sensitive to the length of the mission. A premium should be placed both on the instrument's long-term reliability and on having at least the option of a low-cost operations mode, suitable for an extended mission.

Acknowledgments. The work on POINTS was supported in part by NASA and the Smithsonian Institution, with additional contributions by JPL, Itek, and TRW. The work on Newcomb was supported by the Smithsonian Institution and the USNO.

Appendix. Detection threshold when searching for rare events.

Introduction. There is a class of experiments whose object is to look for a rare event by repeatedly making a measurement, i.e., looking in multiple "channels." Traditionally, a serious problem for such experiments is false alarms. A "3σ event" may be rare, but if 10,000 measurements are made, one expects to see about 13 of them under the assumption that the errors are Gaussian. A further complication comes from the non-Gaussian noise found empirically to corrupt many experiments. This latter complication is not addressed here.

In this Appendix, we consider the question: At what level must we set the detection threshold for the individual measurements in order that the experiment have the required reliability? The answer to this question need not be very precise. Experimental uncertainties are often not known to better than 10%.

Analysis. Let x be a measured quantity which has a zero-mean Gaussian distribution with standard deviation σ. Let $P(B)$ be the probability that x is greater than $B\sigma$ for a given measurement in which the sought-after rare event is absent. We assume that this probability is small. If there are N measurements, then the probability that $x > B\sigma$ in at least one of them is

$$P_N(B) = 1 - [1 - P(B)]^N \approx NP(B)\left[1 - \frac{N-1}{2}P(B) + \ldots\right] \quad . \tag{5}$$

To an excellent approximation, we are free to keep just the first term of the series since we are considering reliable experiments for which $NP(B)$ is small.

It is convenient to talk about an experiment in terms of confidence limits, which are often translated to the number of standard deviations from the mean that a result represents. When this terminology is used, a single-parameter Gaussian distribution is implicit. We may determine A, the number of standard deviations that characterizes the entire experiment, according to

$$P(A) = P_N(B) \approx NP(B) \quad . \tag{6}$$

For a Gaussian probability density function, the cumulative probability function $P(A)$ has an approximation for large A

$$P(A) \approx \frac{1}{A\sqrt{2\pi}}\, e^{-A^2/2}\,(1 - A^{-2} + \ldots) \tag{7}$$

(Gautschi 1968). By combining Eqs. (6) and (7), and dropping small terms, we obtain

$$B^2 \approx A^2 + 2\log_e N + 2\log_e\left(\frac{A}{B}\right) + 2\log_e\frac{1 - B^{-2}}{1 - A^{-2}} \tag{8}$$

We can obtain a first approximation to the solution to Eq. (8), adequate for most purposes when A is not small, by neglecting the terms on the right containing B:

$$B \approx \sqrt{A^2 + 2\ln N} \tag{9}$$

For example, in an experiment with 10,000 measurements, a "3σ result" requires a single-event significance of 5.2σ.

References

Babcock, R.W., Chandler, J.F., Phillips, J.D., and Reasenberg, R.D. 1997, in preparation

Casertano, S. 1996, private communication

Chandler, J.F. & Reasenberg, R.D. 1990, in Inertial Coordinate System on the Sky, J.H. Lieske and V.I. Abalakin, Dordrecht: Kluwer Academic Publisher, 217

Gautschi, W. 1968, in Handbook of Mathematical Functions, M. Abramowitz and I.A. Stegun, National Bureau of Standards Applied Mathematics Series 55, 298

Lattanzi, M.G., Spagna, A., Sozzetti, A., & Casertano, S. 1997, in these proceedings

Noecker, M.C. 1995, in Spaceborne Interferometry II, R.D. Reasenberg, Bellingham: SPIE 2477, 188

Peterson, D. et al. 1996, The Space Interferometry Mission: Taking the Measure of the Universe. Report to NASA Office of Space Science. Available at ftp://ftp.ess.sunysb.edu/pub/siswg/

Phillips, J.D. et al. 1995, in Spaceborne Interferometry II, R.D. Reasenberg, Bellingham: SPIE 2477, 167

Reasenberg, R.D. 1984, BAAS, 16, 758

Reasenberg, R.D. & Shapiro, I.I. 1986, in Relativity in Celestial Mechanics and Astrometry, IAU Symposium No. 114, J. Kovalevsky & V.A. Brumberg, Dordrecht: Reidel, 383

Reasenberg, R.D. et al. 1988, AJ, 96, 1731

Reasenberg, R.D. 1993, in The First William Fairbank Meeting on Relativistic Gravitational Experiments in Space, M. Demianski & C.W.F. Everitt, Singapore: World Scientific, 81

Reasenberg, R.D., Babcock, R.W., Noecker, M.C., & Phillips, J.D. 1993 Remote Sensing Reviews (Guest Editor: J. Alexander), 8, 69

Reasenberg, R.D., Babcock, R.W., Phillips, J.D., Johnston, K.J., & Simon, R.S. 1993, in Spaceborne Interferometry, R.D. Reasenberg, Bellingham: SPIE 1947, 272

Reasenberg, R.D. et al., 1996, in Space Telescopes and Instruments IV, P.Y. Bely & J.B. Breckinridge, Bellingham: SPIE 2807, 32

Shapiro, I.I. 1978, private communication

Working Papers 1991. Astronomy and Astrophysics Panel Reports. National Research Council, National Academy Press, Washington, DC